Journal of African History, **40** (1999), pp. 469–473. Printed in the United Kingdom

© 1999 Cambridge University Press

REVIEW ARTICLE

LINGUISTIC EVIDENCE AND HISTORICAL RECONSTRUCTION

BY JAN VANSINA
University of Wisconsin–Madison

A Green Place, a Good Place : Agrarian Change, Gender and Social Identity in the Great Lakes Region to the Fifteenth Century. By DAVID LEE SCHOENBRUN. Portsmouth, NH: Heinemann: Oxford: James Currey, 1998. Pp. xiv + 301. £40 (ISBN 0-325-00041-7); £15.95, paperback (ISBN 0-325-00040-9).

An African Classical Age : Eastern and Southern Africa in World History, 1000 B.C. to A.D. 400. By CHRISTOPHER EHRET. Charlottesville: University Press of Virginia; Oxford: James Currey, 1998. Pp. xvii + 354. £35 (ISBN 0-8139-1814-6).

Recently several historical reconstructions based on linguistic evidence and dealing with ancient times have been published in African history. In 1998 alone there are the two books reviewed here as well as a major work by Gerda Rossel.[1] Linguistic sources contribute much to the recovery of aspects of the past, which would otherwise remain out of reach, and the standard methodologies of historical linguistics are well known to readers of this journal.[2] Yet in practice many historians remain all too often disconcerted by such studies because they have great difficulty in evaluating them: i.e. in linking assertions made to the evidence provided and so to establish the credibility of such statements. This is not just because many historians are unfamiliar with linguistic evidence but because all the evidence necessary for evaluation is usually not available in the work studied, and often enough authors do not clearly indicate where it can be found. Indeed sometimes it is not available at all. In such cases one has to take the statements made by the authors on faith: one believes the author or not. That is clearly unacceptable. For is it not a fundamental rule in history writing that assertions must be substantiated and hence evidence must be cited or provided? Any work without substantiation cannot be considered to be a work of history at all.

Works based on historical linguistics do present special difficulties in this regard. In many cases one requires data on the historical classification and histories of the languages involved, as well as the lexical data used in their full forms, and with the range of their meanings. If not available elsewhere the inclusion of such data threatens to overwhelm a manuscript and is certainly a reason why few of the doctorates based on linguistic evidence are ever published. A prime example is J. J. Hoover's lengthy and valuable dissertation, more than half of which is devoted to an exposition of the 'background' evidence.[3] Publishers tend to reject such manuscripts at first sight. Yet…historians must produce their evidence. What is one to do?

[1] Gerda Rossel, *Taxonomic–Linguistic Study of Plantain in Africa*, (Leiden, 1998).

[2] Derek Nurse, 'The contributions of linguistics to the study of history in Africa', *Journal of African History*, 38 (1997), 359–91, for the latest overview.

[3] J. J. Hoover, 'The seduction of Ruwej: reconstructing Ruund history (The Nuclear Lunda; Zaire, Angola, Zambia)' (Ph.D. thesis, Yale University, 1978).

https://doi.org/10.1017/S0021853799007598 Published online by Cambridge University Press

Schoenbrun found one solution. He published the bulk of his evidence separately in a long technical article (the genetic language data) and in a dictionary (the words used in comparative perspective).[4] His *Green Place* must be evaluated along with the other two works at hand. These happen to be exemplary. In the article he sets forth exactly what data he used to set up the genealogical tree for the languages of the Great Lakes area which he adopts, and he also details how this result compares to trees proposed by other researchers before him. It is on this basis that he is able to provide the relative chronology adopted in *Green Place*. His dictionary not only presents forms with all their features (correct vowels, tones, prefixes) with their full distribution and all variants of meaning, as well as an overall gloss, but also provides an etymology and tells us when this word first appeared as an innovation or a transfer in terms of levels of language differentiation. With this in hand, *A Green Place* can refer to the evidence when necessary and can concentrate on the historical reconstruction itself.

The book gives us a broad overview of agricultural history in the area from *c.* 1000 B.C. to A.D. 1500 and a more detailed discussion of the history of social identities, politics and health in two chronological slices and three areal ones. Schoenbrun has been well served by what may be the densest and richest ethnographic information for any area comparable in size in Africa, that is a set of data that convey the 'meaning' of the words whose 'form' establishes their place in the linguistic history and chronology. In addition, he also made use of palaeo-environmental and archaeological information. Indeed he is one of the few who informs us about the genuine danger of correlating sites too readily with speakers of a specific language (pp. 28–37, and 46–8).

As a result his reconstruction is both rich and convincing. It is a superb specimen of the kind of results historians can expect from linguistic sources and of their 'resolution': i.e. the degree of detail in substance and chronology which can be detected by these means. For while the findings are many and fairly detailed, this reconstruction cannot be but a history without concrete agents or events, and hence some may feel it never quite comes into sharp focus. On the other hand semantic sources reveal thinking by the actors themselves and lead to crucial insights of which Schoenbrun's distinction between creative and instrumental power is perhaps the most impressive.

The author tells his story well, especially in the last three chapters where he deals with the regional developments from 800 to 1500. There is much here that is novel, unexpected, and surprising and one often agrees with the interpretations made by the author. But even where one does not, one can still appreciate the exact data on which the argument rests and be convinced that they do require an explanation. This implies that this work will be fruitful, because its data base will allow others both to extend reconstructions and in some cases to provide alternative explanations.

The author of *The Classical Age*, Chris Ehret, is a senior historian who, as a researcher and a teacher at UCLA, has devoted his whole career to the reconstruction of history based on linguistics; Schoenbrun was one of his students. Ehret became well known for his reconstruction of various language groups and families connected with Eastern Africa.[5] In all his work he relies on a combination

[4] David Lee Schoenbrun, *The Historical Reconstruction of Great Lakes Cultural Vocabulary: Etymologies and Distributions*, (Cologne, 1997) and 'Great Lakes Bantu: classification and settlement chronology', *Sprache und Geschichte in Afrika*, 14 (1994), 1–62.

[5] *Southern Nilotic History: Linguistic Approaches to the Study of the Past* (Evanston, 1971); *Ethiopians and East Africans: The Problems of Contact* (Nairobi, 1974); *The*

https://doi.org/10.1017/S0021853799007598 Published online by Cambridge University Press

of lexicostatistics and on lexical innovations, including 'loans', for he saw from the outset the particular historical importance of lateral transfers between languages belonging to different linguistic families.[6] His use of lexicostatistics is not simple. He uses his own lists, and adjusts results by estimating corrections for geographical proximity and for perceived loans.[7] The method of lexical innovations requires that a word be found in the languages of a supposed group *and nowhere else*. Negative evidence is crucial to the method and one must therefore have sufficient data not only from the neighboring languages of a supposed group, but also from all potentially related languages – in the Bantu case very many! This technique is accordingly not as straightforward as it looks.[8] To evaluate the weight of any interpretation based on these techniques the reader must therefore have access to all of the evidence and not just to a conclusion derived from the data.

In 1973, Ehret published a history of the settlement of eastern and south-eastern Africa based on linguistics, which is the forerunner and the model for this book.[9] In the latter book he uses, once again, the internal classification of the greatest portion of the East Bantu group of languages, which he calls Mashariki,[10] as the means to construct a history of a 'Greater East Africa', from Kenya to Natal. He divides the time span studied in two major eras: the first one 1000–300 B.C., when East Bantu speakers were still confined to the great Lakes area, the second 300 B.C.–A.D. 400, when they spread out over all of eastern and southern Africa. He devotes four chapters to the first era and three to the second, dealing in each part with migrations, natural surroundings, and economic, social, political, religious and cultural practices and change. The evidence cited for his reconstruction consists of 55 tables and 284 footnotes.

The least that can be said is that this evidence is not transparent. His cavalier attitude towards evidence appears already in his first table (p. 36), entitled 'Generally recognized subgroupings of modern-day languages belonging to the Mashariki group'. He does not present, discuss or refer to linguistic work by others about this question; not all his subgroupings are universally recognized, nor is his nomenclature (Kaskazi, Kusi) accepted. He is in fact the only scholar to class the Nyasa languages (his number 12) under Southeast Bantu. Moreover, in the commentary on this table (p. 35) he refers to his own earlier classification and nomenclature and tells us that the 'reanalysis of the loanword evidence' requires

---

Historical Reconstruction of Southern Cushitic Phonology and Vocabulary (Berlin, 1980); 'Proto-Cushitic Reconstruction', *Sprache und Geschichte in Afrika*, 8 (1987), 7–180; *Reconstructing Proto-Afroasiatic* (*proto-Afrasian*)*: Vowels, Tone, Consonants and Vocabulary* (Berkeley, 1995); *A Comparative Historical Reconstruction of Nilo-Saharan*, in press.

[6] C. Ehret, 'Cattle keeping and milking in eastern and southern African history', *Journal of African History*, 8 (1967), 1–17.

[7] For example, see C. Ehret and M. Kinsman, 'The dialect classification of Shona and its implications for Iron Age history in southern Africa', *International Journal of African Historical Studies*, 14 (1981) 401–43. See also P. Piron, 'Classification interne du groupe bantoïde' (3 vols.) (Ph.D. dissertation, Université Libre Brussels, 1996), 533–61, for a thorough discussion of the greatly superior methodology, including statistical techniques, used today to evaluate the effect of geographic proximity.

[8] See Piron, 'Classification', 539–42, for the advantages and drawbacks of this technique, which one should use only for small groups.

[9] 'Patterns of Bantu and central Sudanic settlement in central and southern Africa (*c.* 1000 B.C.–500 A.D.)', *Transafrican Journal of History*, 3 (1973), 1–71.

[10] 'Mashariki' is a name Ehret coined after Swahili 'east', itself derived from Arabic *mashriq*, to designate a narrow East Bantu which differs from other classifications by excluding the group composed of the languages of Maniema, Luban and most Zambian languages (in standard reference the new group D, L and M).

https://doi.org/10.1017/S0021853799007598 Published online by Cambridge University Press

this redefining. But we are not told what those loanwords are and why they lead to this change! Then, much later in the same chapter, we are given a list of eight subgroups of Kaskazi by 'middle Mashariki times' according to a classification by D. Nurse 'which I have modified in several respects' (p. 55). Nurse had published a 100-page-long study carefully detailing all the evidence he had for setting up his subgroupings.[11] But that counts for nothing when Ehret combines two groups of Nurse into one and then twice splits another group of Nurse into two, all without giving any reason for doing so, except in one instance where he deigns to refer us in brackets to 'Waite and Ehret, unpublished'.

All the other tables consist of starred forms with a gloss. But again the evidence is not transparent. Too much is missing. Thus Table 2 (p. 38) presents 'The natural environment of early Mashariki society' and contains eleven terms. Taking the first subset which includes terms of very wide Bantu or proto-Bantu distribution, we find an entry: '**-sàká*, secondary forest'. Not only does the form, like all other Bantu forms cited in this book, lack the necessary indication of class, but the initial consonant of the root is **s* which does not occur among Proto-Bantu phonemes. The reader is not told that the form occurs in M. Guthrie's *Comparative Bantu* as '**-càká*', glossed as 'thicket (bush country)' and followed by a list of reflexes in various languages and attributed to common Bantu.[12] In the same subset '**-súbí*, bush country, savannah' has similar shortcomings. It also occurs in Guthrie, although with a very incomplete distribution.[13] Ehret, then, has changed that phonemic inventory all by himself and against the consensus of all other linguists, and without a word of explanation unless we accept the cavalier note (p. 28) 'for the general reader' that 'in many cases the detailed phonological reconstructions that validate language evidence appear in other works, and the reader who wishes to assess such data firsthand should turn to those materials'. In this case that search would be in vain. And so would a search for the phonemic inventory of Mashariki or its successors.

Nor does any of Ehret's tables inform us sufficiently about distributions (i.e. languages in which reflexes of the form are found), etymologies or the semantic fields corresponding to each form. And we can only guess where his data come from. Are they his own collections in the field, those of others in manuscript, or published dictionaries and vocabularies? The upshot is that anyone who dares to question any item in these tables has to redo the entire research for themselves. Is this acceptable in works of history?

An examination of the footnotes in the search of further evidence also yields some disquieting results: 284 notes are not many for a work of this scope and size. Most of them cite a single reference and a few do not contain any. In addition, quite a few citations are useless to readers as they refer to inaccessible material such as 'personal communications', 'in preparation', 'never published' and unpublished papers: e.g. see Ehret 1964, 1973a, 1992, or p. 237 n. 35 [not even in the bibliography], Katanekwa 1995 [should be 1994], Waite and Ehret 1981, Werner 1973. There are also references to papers given in his seminar (Laumann 1995, Gonzales 1995) and even to an oral presentation at a conference (p. 207, n. 32)![14]

Most of the citations are to Ehret himself. Some of these lead the reader to further evidence, while others lead only to further unsubstantiated statements.

[11] 'The diachronic background to the language communities of south-western Tanzania', *Sprache und Geschichte in Afrika*, 9 (1988), 15–155.

[12] M. Guthrie, *Comparative Bantu*, iii (Farnborough, 1970), 78, C.S. 260.

[13] *Ibid* 117, ps 124.

[14] This is in contrast to citations to archival material in repositories accessible to other scholars, whereas they do not normally have access to the statements or documents cited here.

https://doi.org/10.1017/S0021853799007598 Published online by Cambridge University Press

Next in number come citations of dissertations he supervised. These are often circular in that they refer back to statements by Ehret, not to evidence.[15] That leaves a small minority of citations to others, none of which cite references critical of Ehret's previous work, or which are at variance with his views here.[16] As an illustration of the situation let us take Chapter 8. Of the 38 footnotes, 36 include references. Of these, 18 are references to himself (of which eight are to as yet unpublished material), six are to his students (of which four are to their dissertations, one is to 'personal communication' and one is to a published article). That leaves 14 for other authors of which one is a 'personal communication' and the rest refer to nine publications by nine different authors.[17] All in all, then, the footnote apparatus fails to provide the necessary evidence.

Moreover, the parochial concentration on his own school is excessive and unhealthy. This is very evident in the index which contains the names of only nine scholars cited in the text: no linguists, five of his own students, three archaeologists, and this reviewer, who often serves as a general scapegoat. Yet at least two linguists, Guthrie and Nurse, are frequently cited in the text. Why are they and others not even in the index? Surely such an attitude, linked to the weakness of the evidentiary base, explains why many readers feel that this author is intolerably arrogant. Readers feel that they must accept the author's statements as dogma, meekly and without question, just as many of Ehret's students appear to have done in their dissertations. There are, then, the faithful who know and unbelievers who do not. In such a climate scholarship becomes impossible.

The citation of evidence, transparent evidence, is the bedrock of all historiography without which assertions cannot be evaluated. The only sensible stance in such cases is to refuse to believe or disbelieve, and to treat the assertions made merely as speculative suggestions. Once evidence is provided in the future many of these statements may well turn out to be correct, but others will certainly not be. As of now, Ehret's book remains merely a narrative set of suggestions. In sharp contrast, the books by Schoenbrun and Rossel give access to the evidence and can be assessed, and a review's claims that they are sound history is an assessment based on an evaluation of that evidence. Indeed, both Schoenbrun and Rossel demonstrate in their different ways just how linguistic data can lead to valid historical reconstructions.

[15] A blatant case of this situation is the referral by the faithful to Ehret's opinions about sub-groupings and Bantu expansion which he himself never substantiated in print. Even in this book (pp. 43, 45–7, esp 46 n. 16, and his promises in n. 17 'forthcoming') he relies on his own judgment of 1964 and other material now twenty years old, while ignoring the results of massive research that has become available since then.

[16] For example, Ehret and Kinsman on Shona, which was refuted by C. H. Borland: 'Conflicting methodologies of Shona dialect classification', *South African Journal of African Languages*, 1 (1984), 1–12, but is still cited (pp. 226, n. 14; 227, n. 19), as if Borland did not exist. It is striking that Ehret never cites Hoover, *Section*, despite its detailed and varied evidence and its sophisticated linguistic treatment of the material.

[17] Or e.g. Chapter 4: 37 notes, 36 with references: 20 to himself, ten to his students, nine to others. But only five are to publications.

https://doi.org/10.1017/S0021853799007598 Published online by Cambridge University Press